

10. Royal College of Physicians. *Retention of Medical Records—with Particular Reference to Medical Genetics* 2nd edn. (Royal College of Physicians Publications, London, 1998).
11. Medical Research Council. *Personal Information in Medical Research*. (Medical Research Council Publications, London, 2000).

12. Medical Research Council. *Human Tissue and Biological Samples for Use in Research, Operational and Ethical Guidelines*. (Medical Research Council Publications, London, 2001).
13. Nuffield Council on Bioethics. *Genetic Screening: Ethical Issues*. (Nuffield Council Publications, London, 1993).

14. Nuffield Council on Bioethics. *Human Tissue: Ethical and Legal Issues*. (Nuffield Council Publications, London, 1995).
15. Human Genome Organization (HUGO) Ethics Committee. *Statement on the Principled Conduct of Genetics Research*. (HUGO International, London, 1996).

## MEDLINE definitions of race and ethnicity and their application to genetic research

To the editor

Over the last five years, the use of MEDLINE has increased more than ten-fold, attesting to the importance of the database in the scientific community (see [http://www.nlm.nih.gov/bsd/medline\\_growth\\_508.html](http://www.nlm.nih.gov/bsd/medline_growth_508.html)). MEDLINE must organize thousands of topics into an indexing system that strikes a balance between chronological consistency and current interests. It is an invaluable resource to scientific research. But race- and ethnic group-related terms in the Medical Subject Headings (MeSH) used to index articles in MEDLINE are decades old and inconsistent. In light of recent debates concerning how best to describe human genetic variation, this valued resource requires attention so that it can continue to support high quality research.<sup>1-3</sup>

A search for 'race' in MeSH returns the phrase "racial stocks", which MeSH defines as "major living subspecies of man differentiated by genetic and physical characteristics" (see <http://www.nlm.nih.gov/mesh/MBrowser.html>). MeSH lists four races: Caucasoid, Mongoloid, Negroid and Australoid. It defines each of the first three as a "major racial group" distinguished "according to physical features." The definition also lists each stock's geographic regions and some of its populations, which for Negroid include "Amdanemane, Afr [sic] Bushmen, Half-Hamites, Hottentots, Melanesians, Negritos, Negritos, Papuans, Pygmies, Semangs" (see <http://www.nlm.nih.gov/mesh/MBrowser.html>). Several of these outdated terms appear in the same 1950 publication that provides the original source of this definition for Australoid (ref. 4 as cited in ref. 5). This definition omits reference to genetic and physical characteristics and lists only the names of some Australoid populations, such as the Veddas of Ceylon. The Ainu of Japan are assigned to both Mongoloid and Caucasoid racial stocks (see <http://www.nlm.nih.gov/mesh/MBrowser.html>).

MeSH defines ethnic group as "a group of people with a common cultural heritage that sets them apart from others in a variety of social relationships." MeSH lists 13 such groups, drawn primarily from United States populations. Ethnic group definitions range from one phrase to several sentences. The characteristics listed include, in order of frequency cited, geographic location (nine), racial classification (six), ancestry (three) and history (two). Three characteristics are used only once: religion to define Jews; social organization ("itinerant life and tribal organization") to define gypsies; and language group (Semitic) to define Arabs. 'Aborigine' is defined as a kind of population, "indigenous inhabitants," rather than a specific population "with a common cultural heritage," contradicting MeSH's own ethnic group definition (see <http://www.nlm.nih.gov/mesh/MBrowser.html>).

MeSH distinguishes between racial stocks and ethnic group. The former is designated for indexing articles that concern a population's "physical characteristics, genetic characteristics, anthropometric measurements, physiology," whereas the latter designates articles about "cultural, psychological, social, sociological or ethnological aspects." MeSH instructs indexers to override an author's choice of race and ethnicity terms if it conflicts with MeSH's definitions (see [http://www.nlm.nih.gov/mesh/indman/chapter\\_30.html](http://www.nlm.nih.gov/mesh/indman/chapter_30.html)). Searches conducted in October 2002 suggest, however, that indexers apply this distinction inconsistently and that in some categories, as many as 30% of articles conforming to racial stock criteria are instead indexed to ethnic group.

Additional searches, also conducted in October 2002, showed little need for several outdated and sometimes offensive<sup>4</sup> terms used in racial and ethnic group definitions. For example, 'Hamites' is listed as a Caucasoid population, but as a keyword search, the term returns only one article<sup>5</sup>.

'Hamitic-Semitic' subjects are referred to in two articles.<sup>6,7</sup> From the Negroid racial stock definition, 'Hottentots' returns a handful of articles, mostly historical. 'Negritos' and 'Half-Hamites' each return the message that MEDLINE is "[u]nable to map your term to a subject heading."

MeSH should conform with standard language and abandon nineteenth century colonialist terms, such as Hottentots, Hamites and Half-Hamites. Furthermore, in addition to a general review and standardization of racial and ethnic group terms, MeSH needs to reconsider its assertion in the definition of racial stocks that there are sub-species among humans. For historical accuracy, MEDLINE might devise a way to identify articles that ascribe to the sub-species argument. But the implication that this long-repudiated notion is current and legitimate is unacceptable.<sup>10,11</sup>

Pamela Sankar

Department of Medical Ethics, School of Medicine, University of Pennsylvania, Philadelphia, Pennsylvania, 19104, USA.  
e-mail: [sankarp@mail.med.upenn.edu](mailto:sankarp@mail.med.upenn.edu)

1. Stolberg, C.G. *New York Times* (New York, 2001).
2. Collins, F.S., & Mansoura, M.K. *Cancer* **91** (1 Suppl), 221-225 (2001).
3. Jorde, L., et al. *Proc. Natl. Acad. Sci. USA* **94**, 3100-3103 (1997).
4. Coon, C., Gorn, S. & Bircselli, J. *Races: A Study of the Problems of Race Formation in Man* (CC Thomas, Springfield, 1950).
5. Molnar, S. *Races, Types, and Ethnic Groups: The Problem of Ethnic Variation* (Prentice-Hall, Englewood Cliffs, New Jersey, 1975).
6. *The American Heritage Dictionary of the English Language* (ed. Soukharov, A.) (Houghton Mifflin, Boston, 1996).
7. Arnaiz-Villena, A., Martinez-Laso, J. & Alonso-Garcia, J. Iberia: population genetics, anthropology, and linguistics. *Human Biology* **71**, 725-43 (1999).
8. Armstrong, J.C. *Trans. R. Soc. Trop. Med. Hyg.* **72**, 342-344 (1978).
9. Mathews, H.M. & Armstrong, J.C. *Am. J. Trop. Med. Hyg.* **30**, 299-303 (1981).
10. Sankar, P. & Cho, M. *Science* **298**, 1337-1338 (2002).
11. Braun, L. *Perspect. Biol. Med.* **45**, 159-174 (2002).

